Advanced Microeconomics

Lecture 8: Information Design and Rubinstein Bargaining

Yihu Hou

Renmin University of China

Spring 2026

Outline

1. From mechanism design to information design
2. Bayesian persuasion: why partial disclosure can beat full disclosure
3. General framework: posterior beliefs, Bayes plausibility, and concavification
4. Rubinstein bargaining: discounting, first-mover advantage, and bargaining power

From Lecture 7 to Lecture 8

Lecture 7: mechanism design

The designer controlled:

- allocation rules,
- payment rules,
- IC and IR constraints,
- who gets what and who pays what.

Lecture 8: information design

The designer now controls:

- what is disclosed,
- how noisy the disclosure is,
- when information arrives,
- which beliefs the receiver forms before acting.

Main question

How should a principal design a mechanism when agents privately know their types?

Main question

How should a sender or information designer choose an experiment so that the receiver takes a more favorable action?

Common backbone

Both lectures study **strategic response under private information**. Lecture 7 changed **menus and transfers**; Lecture 8 changes **beliefs**.

What an information designer actually commits to

Persuasion timeline

Key object

An experiment is a mapping

$$\sigma(m \mid \theta),$$

the probability of message m in state θ.

What is chosen ex ante?

- The sender does *not* promise a specific message.
- The sender commits to a **rule for generating messages**, possibly from data the sender does not observe ex post.
- After seeing the message, the receiver updates beliefs and chooses a best response.

Interpretation

The commitment is to an **information structure**: a truthful rule for generating messages, not a lie.

Persuasion is not cheap talk

Cheap talk

- messages are costless and unverifiable,
- the sender can simply say anything,
- credibility comes only from equilibrium incentives.

Bayesian persuasion

- the sender commits to a signal experiment,
- messages are generated according to that committed rule,
- credibility comes from commitment to the experiment.

One-sentence distinction

Cheap talk asks whether a message is **believable**. Persuasion asks how to design a **truthful but possibly coarse** information structure.

A binary persuasion environment

State and action

There are two states:

$$\theta \in \{H, L\}, \qquad \Pr(H) = \lambda.$$

The receiver chooses

$$a \in \{0, 1\}.$$

Interpret $a = 1$ as buy / invest / approve, and $a = 0$ as reject.

Payoffs

Receiver:

$$u_R(H, 1) = 1, \qquad u_R(L, 1) = -1, \qquad u_R(\theta, 0) = 0.$$

Sender:

$$u_S(\theta, 1) = 1, \qquad u_S(\theta, 0) = 0.$$

Conflict of interest

The receiver wants action 1 only when the posterior chance of H is high enough. The sender wants action 1 as often as possible.

The receiver uses a posterior threshold rule

Expected payoff from action 1

If the posterior is

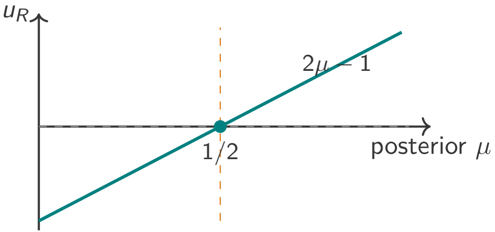
$$\mu = \Pr(H \mid m),$$

then the receiver's payoff from taking action 1 is

$$\mu \cdot 1 + (1 - \mu) \cdot (-1) = 2\mu - 1.$$

So the receiver chooses $a = 1$ iff

$$\mu \geq \frac{1}{2}.$$



Meaning

The question is not "what is the true state?" but "is the posterior high enough to justify action 1?"

Two naive benchmarks: no disclosure and full disclosure

No disclosure

If the sender reveals nothing, the receiver acts on the prior λ:

$$a = \begin{cases} 1, & \lambda \geq \frac{1}{2}, \\ 0, & \lambda < \frac{1}{2}. \end{cases}$$

So sender payoff is

$$V^{ND}(\lambda) = \begin{cases} 1, & \lambda \geq \frac{1}{2}, \\ 0, & \lambda < \frac{1}{2}. \end{cases}$$

Full disclosure

If the sender reveals the state perfectly, the receiver takes action 1 only in state H. So sender payoff is

$$V^{FD}(\lambda) = \lambda.$$

Immediate comparison

If $\lambda < 1/2$, then full disclosure beats no disclosure because at least the good state still induces action 1. If $\lambda > 1/2$, no disclosure beats full disclosure because it keeps action 1 even in the bad state.

Why partial disclosure can beat both benchmarks

Assume $\lambda < 1/2$. Then no disclosure leads to rejection, while full disclosure yields sender payoff λ.

Idea of partial disclosure

Instead of fully separating H from L, the sender creates two posterior beliefs:

$$\mu_G = \frac{1}{2}, \qquad \mu_B = 0.$$

After the good message, the receiver is just willing to take action 1. After the bad message, the receiver rejects.

Bayes plausibility

Let α be the probability of the good message. The average posterior must equal the prior:

$$\alpha \cdot \frac{1}{2} + (1 - \alpha) \cdot 0 = \lambda.$$

Hence

$$\alpha = 2\lambda.$$

So sender payoff becomes

$$V^P(\lambda) = 2\lambda.$$

Why this is better

For every $\lambda \in (0, 1/2)$, we have $2\lambda > \lambda$. Pooling some bad states with good states makes action 1 occur more often than under full disclosure.

A concrete signal that generates the threshold posterior

Signal rule

Use two messages, g and b:

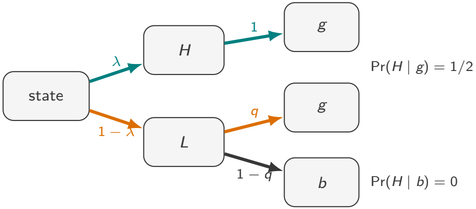
- in state H, always send g;
- in state L, send g with probability q and b with probability $1 - q$.

The posterior after g is

$$\Pr(H \mid g) = \frac{\lambda}{\lambda + (1-\lambda)q}.$$

Set this equal to $1/2$, which gives

$$q = \frac{\lambda}{1-\lambda}.$$



Reading the rule

Good states always pass; bad states pass only often enough to make $\Pr(H \mid g) = 1/2$.

Why the optimal good posterior sits exactly at the threshold

If the good posterior were above $1/2$

The receiver would still take action 1, but the sender could mix in a few more bad states, lower that posterior toward $1/2$, and make the good message occur more often.

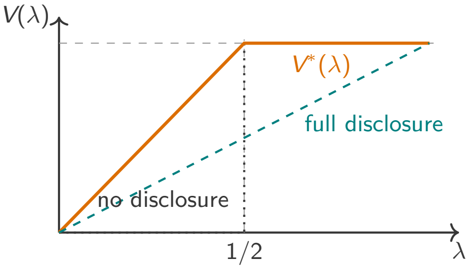
If the good posterior were below $1/2$

The receiver would reject, so that posterior would not help the sender at all.

Knife-edge logic

The sender wants posterior beliefs to be **just persuasive enough**. With a binary action, the optimal informative posterior often lands exactly on the receiver's action threshold.

The sender's value as a function of the prior



Optimal value

$$V^*(\lambda) = \begin{cases} 2\lambda, & \lambda \leq \frac{1}{2}, \\ 1, & \lambda \geq \frac{1}{2}. \end{cases}$$

Interpretation

When the prior is low, persuasion improves on full disclosure by selectively hiding bad news. When the prior is already high, the sender prefers not to reveal anything.

General Bayesian persuasion environment

Primitives

- state $\theta \in \Theta$ with prior π,
- sender commits to an experiment $\sigma(m \mid \theta)$,
- receiver observes message m,
- receiver chooses action $a \in A$,
- payoffs are $u_S(\theta, a)$ and $u_R(\theta, a)$.

Sequence

$$\text{commit } \sigma \longrightarrow \theta \longrightarrow m \longrightarrow a(m).$$

After seeing m, the receiver forms the posterior belief

$$\mu(\theta \mid m)$$

and chooses a best response.

Design problem

Choose the experiment to maximize

$$\mathbb{E}_{\theta,m}\left[u_S(\theta, a(m))\right]$$

subject to receiver best response after every message.

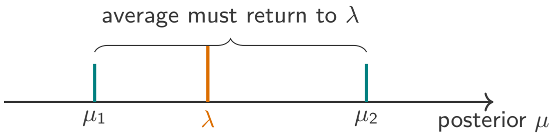
Bayes plausibility: the average posterior must equal the prior

Binary-state version

If the signal induces posteriors $\mu_1, \ldots, \mu_K$ with probabilities $\alpha_1, \ldots, \alpha_K$, then

$$\sum_{k=1}^{K} \alpha_k \mu_k = \lambda, \qquad \sum_{k=1}^{K} \alpha_k = 1.$$

This is why the sender cannot simply choose favorable beliefs freely.



General version

In vector form,

$$\mathbb{E}[\mu(\cdot \mid m)] = \pi.$$

This is the key feasibility constraint in information design.

Reduced form: choose posteriors, not raw messages

Why this simplification works

For the receiver, only the posterior matters. Two different messages that induce the same posterior lead to the same best response.

So the sender can think in two steps

1. Choose a distribution over posterior beliefs.
2. Let the receiver best respond to each posterior.

What remains constrained

The sender still cannot choose arbitrary posterior distributions. They must satisfy Bayes plausibility:

$$\mathbb{E}[\mu] = \pi.$$

Analogy to mechanism design

This is closer to a revelation-principle reduction. We replace raw messages by the posteriors they induce, and Bayes plausibility is the implementability condition for that reduced form.

The sender's interim value as a function of the posterior

Receiver best response first

Given posterior μ, the receiver chooses

$$a^*(\mu) \in \text{argmax}_{a \in A} \mathbb{E}[u_R(\theta, a) \mid \mu].$$

Sender value induced by that response

Define

$$v(\mu) = \mathbb{E}[u_S(\theta, a^*(\mu)) \mid \mu].$$

This tells us how much the sender likes a particular posterior belief.

Binary example

For the earlier example,

$$v(\mu) = \begin{cases} 0, & \mu < 1/2, \\ 1, & \mu \geq 1/2. \end{cases}$$

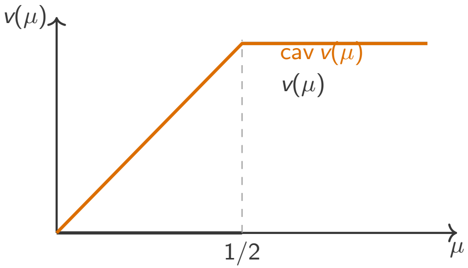
So the sender likes posterior beliefs only because they trigger different receiver actions.

Main problem

Choose a distribution over μ with mean π to maximize

$$\mathbb{E}[v(\mu)].$$

Concavification: the main geometric intuition



Idea

The sender starts from the raw value function $v(\mu)$ and then asks:

What is the highest average sender value achievable by splitting beliefs while keeping the same mean?

The answer is the **concave envelope** cav v.

Practical meaning

Concavification turns persuasion into geometry: replace v by the smallest concave function above it.

Why only a few posterior beliefs typically matter

Binary-state case

The optimal straight line segment touches the graph at only a small number of posterior beliefs. In our binary example, it touches at

$$\mu = 0 \qquad \text{and} \qquad \mu = 1/2.$$

So the optimal signal uses only two posteriors.

General intuition

Optimal persuasion usually does *not* need a very rich message space. It often suffices to use a small set of posterior beliefs that support the relevant segment of the concave envelope.

Teaching intuition

The sender does not want "more information" in general. The sender wants just enough posteriors to move the receiver across the relevant action thresholds.

Where information design appears in practice

Signal precision

- longer free trials,
- more accurate diagnostic tests,
- forecasts with narrower confidence intervals.

Aggregation / coarsening

- average rating instead of all reviews,
- pass/fail grades instead of scores,
- risk categories instead of full distributions.

Timing

- refund windows,
- delayed inspection,
- disclosure before or after a commitment decision.

Economic lesson

Information is another strategic instrument. A designer can change how accurate signals are, how evidence is summarized, or when signals arrive.

Why Rubinstein bargaining matters

One-shot bargaining is too coarse

In a take-it-or-leave-it game, bargaining power looks extreme:

- the proposer makes one offer,
- the responder accepts or rejects,
- there is little room to model delay, patience, or counteroffers.

What Rubinstein adds

Alternating offers let us model:

- delay as a real cost,
- the value of moving first,
- the effect of patience,
- how outside options shift bargaining power.

Core question

If agreement can be postponed and future payoffs are discounted, who gets how much of the surplus?

Rubinstein's alternating-offer model

Environment

- Two players divide a unit surplus.
- At $t = 0$, player 1 proposes a split.
- If player 2 rejects, then at $t = 1$ player 2 proposes.
- This continues forever until agreement.

Discounting

If agreement is reached in period t, player i values one unit received then as

$$\delta_i^t, \qquad \delta_i \in (0, 1).$$

So delay is costly.

Interpretation

The discount factor captures impatience, financing cost, political urgency, legal delay, or deterioration of the trading opportunity.

A two-period warm-up

Last period logic

Suppose there are only two periods. If player 2 gets to propose in period 1, then player 2 can offer player 1 just enough to make acceptance worthwhile. In the limiting benchmark, player 2 keeps almost all of the pie.

What player 2 can guarantee in period 1

If player 2 rejects at $t = 0$, then player 2 becomes the proposer next period. So player 2's continuation value is approximately

$$\delta_2.$$

Acceptance constraint at $t = 0$

Therefore player 2 accepts player 1's first offer only if player 2 receives at least

$$\delta_2$$

in period-0 value.

Warm-up solution

Player 1 offers player 2 exactly δ_2 and keeps $1 - \delta_2$. The responder's reservation value is not zero; it is the discounted value of becoming tomorrow's proposer.

Stationary logic in the infinite-horizon game

Notation

Let x be player 1's share in player 1's current offer.
Let y be player 1's continuation share after one rejection.
Stationarity means the same continuation values recur whenever the same kind of node is reached.

Acceptance constraints and tightness

The responders accept only if

$$1 - x \geq \delta_2(1 - y), \qquad y \geq \delta_1 x.$$

Each proposer wants to keep as much as possible, so both constraints bind in equilibrium:

$$1 - x = \delta_2(1 - y), \qquad y = \delta_1 x.$$

Why tightness is natural

A slack acceptance constraint means the proposer is giving away more than needed. Equilibrium offers leave the responder just indifferent between accepting now and waiting one period.

Solving the bargaining fixed point

Solve for player 1's opening share

Use the binding continuation condition

$$y = \delta_1 x$$

inside

$$1 - x = \delta_2(1 - y).$$

Then

$$1 - x = \delta_2(1 - \delta_1 x)$$

which implies

$$x = \frac{1 - \delta_2}{1 - \delta_1\delta_2}.$$

Player 2's share in player 1's offer

So player 2 gets

$$1 - x = \frac{\delta_2(1 - \delta_1)}{1 - \delta_1\delta_2}.$$

The stationary continuation share for player 1 after a rejection is

$$y = \delta_1 x = \frac{\delta_1(1 - \delta_2)}{1 - \delta_1\delta_2}.$$

Equilibrium structure

Agreement happens immediately. Delay is costly, so each proposer offers exactly the responder's continuation value.

Common discount factor: the textbook formula

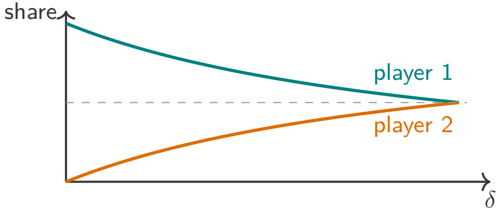
If $\delta_1 = \delta_2 = \delta$

The first proposer's share becomes

$$x_1^* = \frac{1-\delta}{1-\delta^2} = \frac{1}{1+\delta}.$$

The responder gets

$$x_2^* = \frac{\delta}{1+\delta}.$$



Comparative statics

When δ rises, both players become more patient and the first-mover advantage shrinks. As $\delta \to 1$, the split converges to 1/2-1/2.

How to read the formula economically

First-mover advantage

With common $\delta < 1$,

$$\frac{1}{1+\delta} > \frac{1}{2}.$$

Moving first matters because the responder dislikes delay.

Patience

A higher δ_i means player i is more willing to wait. That improves player i's continuation value and therefore their bargaining position.

Why agreement is immediate

The equilibrium already gives the responder exactly the discounted continuation value. Delaying would burn surplus without improving anyone's terms.

Economic lesson

Bargaining power does not come from rhetoric here. It comes from **what happens if I say no and wait one more period**.

Unequal patience means unequal bargaining power

Player 1's opening share

$$x_1^* = \frac{1-\delta_2}{1-\delta_1\delta_2}.$$

For interior values, this increases with δ_1 and decreases with δ_2.

Interpretation

- A higher δ_1 raises player 1's continuation value after a rejection.
- A higher δ_2 raises player 2's reservation value against player 1's opening offer.

Extreme cases

- If $\delta_1 \to 1$ with $\delta_2 < 1$ fixed, then $x_1^* \to 1$.
- If $\delta_2 \to 1$ with $\delta_1 < 1$ fixed, then $x_1^* \to 0$.
- If both approach 1, write $\varepsilon_i = 1 - \delta_i$. Then $x_1^* \approx \varepsilon_2/(\varepsilon_1 + \varepsilon_2)$, so the limit depends on relative speeds.

Patience = bargaining strength

The player with the lower cost of delay has the stronger outside option inside the bargaining game itself.

Outside options work through the acceptance constraint

Suppose player 2 can exit

If player 2 rejects player 1's offer, player 2 may either

- continue bargaining, worth $\delta_2(1-y)$, or
- exit permanently for outside-option value d.

Modified acceptance condition

Player 2 now accepts only if

$$1 - x \geq \max\{\delta_2(1-y),\ d\}.$$

So player 1 must leave player 2 at least that much.

Key point

The outside option may never be exercised in equilibrium, but it still changes the equilibrium split because it raises the responder's reservation value.

Why outside options matter in real economic settings

Labor bargaining

Strike funds, unemployment insurance, or alternative job offers improve the worker's fallback position.

Supply chains

A supplier with alternative buyers can reject a platform's terms more credibly.

Policy example

If farmers know that rejecting a buyer today still leaves a government support price or compensation floor, that fallback may never be used, but it can still improve the negotiated price.

Unifying message

Bargaining power depends on what each side can credibly do after walking away or waiting.

Two separate modules in Lecture 8

Information design

- Static belief-design problem.
- Designer chooses a signal experiment.
- Feasibility comes from Bayes plausibility.
- Main tool: posteriors and concavification.

Rubinstein bargaining

- Dynamic surplus-division problem.
- Players make alternating offers.
- Feasibility comes from acceptance constraints.
- Main tool: continuation values.

End of Lecture 8

- Bayesian persuasion studies how a sender or information designer commits to a signal experiment so that the receiver forms more favorable posterior beliefs.
- The sender cannot choose beliefs freely: posterior beliefs must satisfy Bayes plausibility, meaning their average must equal the prior.
- In the binary example, partial disclosure beats full disclosure because the sender pools some bad states with good states and pushes posterior beliefs just up to the receiver's action threshold.
- The general persuasion problem can be read geometrically through the concavification of the sender's value function over posteriors.
- In Rubinstein bargaining, equilibrium offers are pinned down by each responder's discounted continuation value.
- Patience, first-mover position, and outside options determine bargaining power because they determine what each side can credibly demand after saying no.

Bridge to Lecture 9

What Lecture 8 emphasized

- beliefs can be designed,
- bargaining power comes from timing and outside options,
- strategic outcomes depend on the informational and dynamic environment, not only on primitive payoffs.

What comes next

Lecture 9 turns to IO extensions:

- search frictions and price dispersion,
- Hotelling spatial competition,
- vertical restraints and foreclosure.

Continuity

The next step is still about strategic interaction under frictions, but the friction shifts from information and bargaining delay to search costs, product differentiation, and vertical structure.

References

Selected references for Lecture 8

- Kamenica, E., and Gentzkow, M. (2011). "Bayesian Persuasion." *American Economic Review*.
- Bergemann, D., and Morris, S. (2019). "Information Design: A Unified Perspective." *Journal of Economic Literature*.
- Gentzkow, M., and Kamenica, E. (2016). "A Rothschild-Stiglitz Approach to Bayesian Persuasion." *American Economic Review Papers and Proceedings*.
- Rubinstein, A. (1982). "Perfect Equilibrium in a Bargaining Model." *Econometrica*.
- Osborne, M. J., and Rubinstein, A. (1990). *Bargaining and Markets*. Academic Press.
- Osborne, M. J., and Rubinstein, A. (1994). *A Course in Game Theory*. MIT Press, esp. bargaining chapters.